

## Research Article

# Evaluation Method of IT English Blended Teaching Quality Based on the Data Mining Algorithm

**Lin Shao** 

*Dalian University of Foreign Languages, Dalian, Liaoning 116044, China*

Correspondence should be addressed to Lin Shao; shaolin@dlufl.edu.cn

Received 24 November 2021; Revised 14 December 2021; Accepted 17 December 2021; Published 31 December 2021

Academic Editor: Naeem Jan

Copyright © 2021 Lin Shao. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As the mixed education model gradually becomes widespread in various universities in Japan, the evaluation of the quality of IT English mixed education has become a very important issue, and it is worth considering the corresponding evaluation method. In this paper, we use a data mining algorithm to implement an evaluation method for the interconversion of quantitative data and qualitative concepts and use the IT English mixed teaching model to evaluate and analyze the teaching quality of the course. The evaluation method is feasible and provides a mixing method. Evaluation of the quality of education. Reference method.

## 1. Introduction

With the progress of educational reform in universities, the IT English mixed education model is attracting more and more attention from various universities, and many university teachers conduct practical education and topic research aimed at reforming mixed curriculum education [1, 2]. Blended teaching is becoming more and more important, which is welcomed by the majority of teachers and students. Therefore, what is the quality of education in the IT English mixed education model and how to evaluate the quality of education in this education model are the important issues faced by domestic universities in monitoring the quality of education [3–5]. This requires the establishment of an effective and comprehensive evaluation system for the IT English mixed education model, and we are actively considering various educational quality evaluation methods under the mixed education model. At universities, there are many reference methods for daily class quality evaluation, and from the perspective of fuzzy data processing, there are rough set theory and fuzzy evaluation methods, but they have drawbacks. For example, the rough set theory adopts the idea of removing redundant data, so the reduced information is relatively incomplete. The fuzzy evaluation method has a strong subjectivity in acquiring the fuzzy evaluation matrix, so the rigor of the

evaluation result is low. Due to the randomness and ambiguity of the educational evaluation itself, it is necessary for our evaluation method to be able to reflect the randomness and ambiguity of this evaluation so that the evaluation results are more consistent with the actual situation. In addressing these issues, Thoma et al. used a reverse cloud generator to achieve qualitative and quantitative transformation of evaluation reviews and a virtual cloud algorithm to obtain a comprehensive evaluation cloud model [6–10]. Ye Yong proposed a cloud model-based classroom quality data mining method to estimate the nonlinear relationship between classroom quality and metrics. With the advent of data mining algorithms, data mining algorithms have been applied to various qualitative evaluations of the system and achieved specific results. The application of data mining algorithms in qualitative evaluations can accurately reflect the actual situation of the evaluated object and reflect the randomness and ambiguity of the evaluation itself. This article explores its application in the quality assessment of mixed education using the relevant theory of data mining algorithms. Section 2 of this paper describes data mining. On the basis of this data mining, Section 3 proposes a model evaluation method. Section 4 presents the design of the proposed model. In Section 5, the experimental design and experimental analysis are carried out. In Section 6, the conclusion is given.

## 2. Description of Data Mining

Data mining generally refers to the process of finding hidden information in large amounts of data using a variety of methods such as statistics, online analytics, and intelligence searches. Education data mining refers to the application of multiple disciplines of theory and methods for describing data so that people can quickly distinguish data. Through data mining processes such as clustering and associations, we provide students with video viewing tools for education, research, and solving time problems. For example, teachers can find student learning rules based on student learning data and teach them to effectively improve their teaching methods. This improves educational efficiency and builds effective explanatory models to predict and analyze future education, as shown in Figure 1.

Theoretical methods of data mining models are used to teach evaluation. It mainly uses the three digital feature values of the data mining model (expectations  $E_x$ , entropy  $E_n$ , and superentropy  $H_c$ ) to characterize the results of the evaluation and analyze more detailed information to teach the evaluation. In expectations,  $E_x$  represents the center point of the evaluation value which indicates the definite value of the comment, where expectations  $E_x$  value falls within which comment value range of the evaluation result belongs to the comment level. Entropy  $E_n$  reflects the ambiguity of the comment result,  $E_n$  bigger explains that the comment level is more blurred, superentropy  $H_c$  reflects the degree of dispersion of the evaluation result from the evaluation curve,  $E_n$  explains that the uncertainty of the evaluation result is higher. The three digital feature values of the data mining model can be calculated using the data mining model reverse cloud generator method; that is, the digital feature values of the data mining model for each rating index are obtained from a large number of quantitative ratings. The specific calculation process is as follows: you can get the calculation steps of the reference.

## 3. Model Evaluation Method Based on Data Mining

Data mining models can achieve effective transformations between quantitative and qualitative concepts, thus transforming large amounts of specific evaluation data into corresponding natural comments. According to the evaluated index data, the inverse cloud generator of the data mining model is used to calculate the digital feature values of the data mining model of each index and analyze the meaning of the digital feature values [11, 12]. Because IT English mixed education involves independent learning of online courses and classroom education, the entire teaching process includes many online and offline links. Therefore, in addition to assessing the quality of teachers' traditional classroom education, students should be assessed online. The status of self-learning and the quality of online learning resources are assessed. Due to limited space, this article evaluates and analyzes the learning situation of the school's open course "College English" from the student's perspective and describes the process of applying a data mining model in

educational evaluation. The specific process for applying for an educational evaluation is as follows [13, 14].

*3.1. Specific Evaluation Indicators for Curriculum Design.* In order to better reflect the actual situation of the course education and improve the quality of the course education evaluation, it is necessary to combine the educational practices of specific courses to develop the evaluation index corresponding to the course. If the course is a theoretical course or a hands-on course with a high percentage of hands-on operation, the corresponding metrics should be different to emphasize the characteristics of the course [15]. For example, the training direction of a university's open English course is to pay more attention to the student's language proficiency, and the corresponding content should be reflected in the designed metrics. In addition, the design of the metric system not only considers the comprehensiveness of the metric but also analyzes the importance of each metric in the overall metrics and reasonable weights each metric [16]. Two grades of metrics are set according to the teaching method of the "College English" course. One-step indicators are set up for online learning, learning resources, and classroom learning. Second-level indicators of "online learning" can be set as a result of self-learning attitudes, video viewing progress, and online testing. Second-level indicators under "learning resources" can be set as course video resources, course exercise resources, and course experiment resources. Second-level indicators under "learning in the classroom" can be set as teacher teaching attitudes, teacher teaching abilities, and module test scores [17, 18].

*3.2. Scoring Various Indicators.* Surveys are designed and published, students and related personnel are organized to score various metrics, these scoring data are collected, and some invalid data are processed according to the metrics set summarizing valid data.

*3.3. Calculating the Digital Feature Value of the Data Mining Model of Various Indicators.* Based on the scoring data for each metric, the data mining model reverse cloud generator is used to calculate the data for each metric to get the low-level digital feature values for the data mining model for each metric. These numerical characteristic values reflect the score distribution of each particular indicator.

*3.4. Calculating the Comprehensive Evaluation.* According to the digital function values of the cloud model of the low-level indicator, the data mining calculation method can be used to calculate the data mining model parameters of the high-level indicator layer by layer, and finally, the comprehensive evaluation parameters can be calculated and obtained. The overall assessment is used as the final outcome of this assessment and as a reference for measuring the comment level of the quality of education assessment [19].

*3.5. Set Comment Collection.* According to the practice of daily education evaluation of the school, the final comment

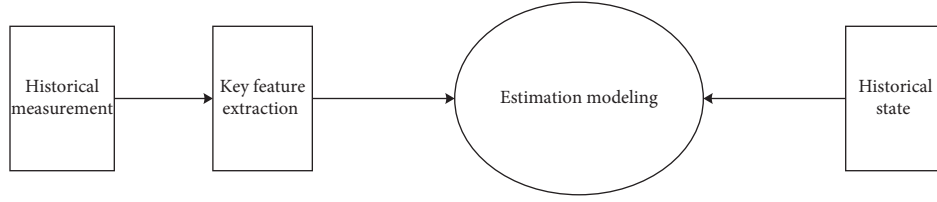


FIGURE 1: Modeling estimation driven by data.

set of the evaluation of the quality of education is set to four levels of “excellent, good, medium, and bad,” and the division interval corresponding to each evaluation level is set and increased. The corresponding “excellent” level score is 9–10 points, the “good” level is 8–9 points, the “medium” level is 6–8 points, and the corresponding score is “bad.” The level is 6–8 points. Values are 0–6 minutes. If the expected value of the overall rating is within the score range corresponding to the “excellent” level, the result of this rating is “excellent.”

3.6. *Drawing a Cloud Map.* Rooting and rendering are simulated and analyzed by MATLAB, etc. The method is similar to the observation point and distribution shape, intuitive judgment, combined conclusion result, etc.

3.7. *Analysis and Evaluation Results.* As comprehensive evaluation combines the cloud digital feature values and the cloud drop distribution state of the cloud map, the evaluation results are analyzed, and more detailed information on educational evaluation is unearthed. Using the cloud model evaluation method, the digital characteristic of each index in each layer of the index can be analyzed concretely, and the evaluation result of each index in the evaluation can be obtained. If the expected value of a certain magical index is low, further improvement is needed in the future, and educational reform can be carried out appropriately. A low module test score in the classroom indicates that the ability to apply knowledge comprehensively is inadequate and that comprehensive content training can be conducted during normal hours. If the expected value of a certain magical index is too high, it means that the magical index is highly recognized. For example, the test shows that the expected value of the teacher’s teaching attitude score is the highest of all indicators, indicating that the student is aware of the teacher’s teaching attitude. If the superentropy value of the indicator is too large, it means that the metric’s score has high volatility, which means that the gap between most scores is large.

## 4. Model Design

4.1. *Dynamic Time Warping Algorithm.* Data processing and analysis must rely on education. Due to different educational situations, the particle size of the collection time for different mixed education data is different. This causes the time dimension to accumulate different data formats, making it impossible to unify the time period. In this scenario, it is

difficult for a typical data mining model to deliver its own performance, so this white paper presents a dynamic time warping (DTW) algorithm for data processing and analysis. The basic principles are as follows:

Suppose the test dataset is  $R$ , the training dataset is  $T$ , and the dimensions of each dataset sample are  $m$  and  $n$ . For supervised learning algorithms, we need to compare the similarities between the test and training datasets. At this time, the similarity of the samples can be measured by calculating the Euclidean distance  $D$  between the samples. If  $n = m$ ,

$$D[T, R] = \sqrt{\sum_{i=1}^M (D[T(i), R(i)])^2}, \quad (1)$$

when  $n \neq m$  needs to introduce the idea of dynamic programming to calculate  $D$ .

As shown in Figure 2, the test sample serial numbers are marked on the  $x$ -axis of the Cartesian coordinate system, and the training samples are marked on the  $y$ -axis. At this time, a cross grid can be formed on the coordinate axes, and the intersection of the grid becomes the intersection of the test sample and the training sample. At the same time, you need to find a path from the lower-left corner to the upper right corner of the axes. The  $i$ -th point that passes through this path is written as  $(n_i, m_i)$ , and the path function is written. That is why

$$\begin{aligned} m_i &= \varphi(n_i), n = i, \\ \varphi(1) &= 1, \varphi(N) = M. \end{aligned} \quad (2)$$

When searching for a path, you need to limit the gradient to ensure the direction of the path. The slope usually varies from 0.5 to 2. If the point of the path that passes through at this time is  $n_{i-1}, m_{i-1}$ , then all possible situations at the next point are

$$\begin{aligned} (n_i, m_i) &= (n_{i-1} + 1, m_{i-1} + 2), \\ (n_i, m_i) &= (n_{i-1} + 1, m_{i-1} + 1), \\ (n_i, m_i) &= (n_{i-1} + 1, m_{i-1}). \end{aligned} \quad (3)$$

At this time, the above equation is used as the constraint condition, the best path in equation (1) is solved, the shortest path is taken as the optimization objective, and the objective function is obtained as follows:

$$\sum_{\substack{n_i=1 \\ m_i=F(n_i)h}}^N D(n_i, m_i) = \min_{F(n_i)} \sum_{n_i=1}^N D(n_i, m_i). \quad (4)$$

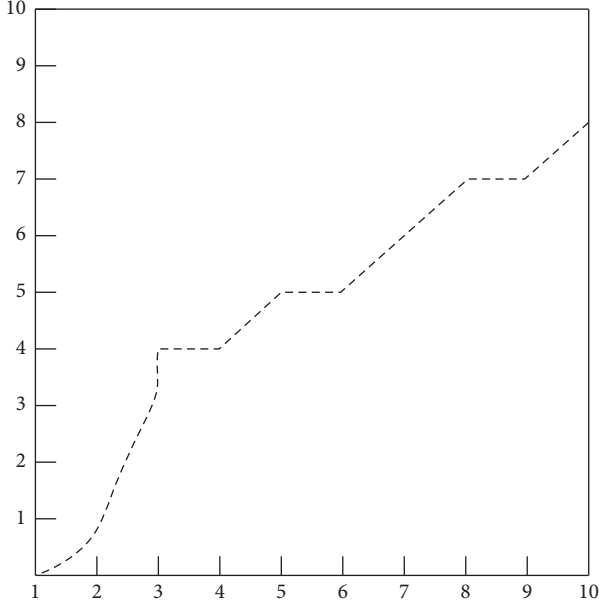


FIGURE 2: DTW algorithm path search principle.

At any point on the axes, only one path can eventually be passed during the path search process. So, for  $(n_i, m_i)$ , there are only three cases  $(n_{i-1}, m_i)$ ,  $(n_{i-1}, m_{i-1})$ , or  $(n_{i-1}, m_{i-2})$  at the previous point on the path. At this time, you need to select the point with the shortest distance from point  $(n_i, m_i)$  to the two points in these three cases as the previous node on the path. At this point, the total distance of the path can be obtained as follows:

$$D[(n_i, m_i)] = \min \left\{ \begin{array}{l} D[(n_{i-1}, m_i)] \\ D[(n_{i-1}, m_{i-1})] \\ D[(n_{i-1}, m_{i-2})] \end{array} \right\}. \quad (5)$$

#### 4.2. Improved DTW Algorithm Based on Dynamic Bending.

In the above equation, the DTW algorithm limits the slope of the path bend during the path search process. However, there are some points where this limitation cannot be met in the actual iterative process. Taking Figure 3 as an example, if the diamonds in Figure 3 perform distance matching, it is not necessary to calculate the distance outside the grid points. Due to the geometric properties of the diamond, it is not necessary to store the cumulative distance of every step in the calculation.

In Figure 3, the actual bend is divided into three path segments:  $(1, X_a)$ ,  $(X_a + 1, X_b)$ , and  $(X_b + 1, N)$ , and the coordinates have the following relationship:

$$\begin{cases} X_a = \frac{1}{3}(2M - N), \\ X_b = \frac{2}{3}(2N - M). \end{cases} \quad (6)$$

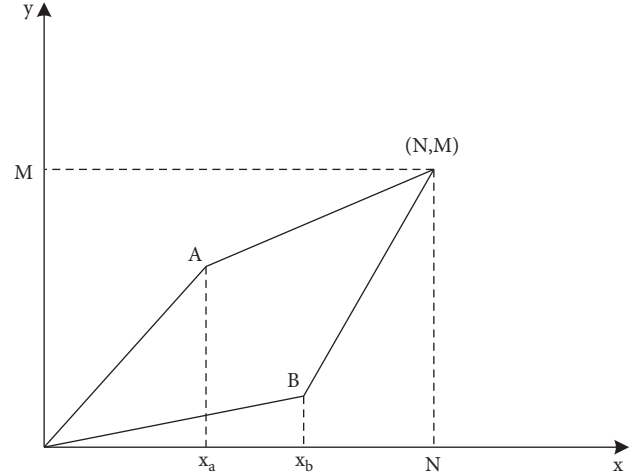


FIGURE 3: Schematic diagram of algorithm path constraint.

Taking  $X_a$  and  $X_b$  as close integers, the constraint conditions for dynamic bending matching can be obtained at this time:

$$\begin{cases} 2M - N \geq 3, \\ 2N - M \geq 2. \end{cases} \quad (7)$$

After introducing dynamic bending, there is no need to compare the X-axis feature vector with the feature vector corresponding to the Y-axis point. All you have to do is to compare it to the eigenvectors of  $[y_{\min}, y_{\max}]$ . The calculation method of the endpoints of this section is as follows:

$$y_{\min} = \begin{cases} \frac{1}{2}x, 0 \leq x \leq X_b, \\ 2x + (M - 2N), X_b \leq x \leq N, \end{cases} \quad (8)$$

$$y_{\max} = \begin{cases} 2x, 0 \leq x \leq X_a, \\ \frac{1}{2}x + (M - \frac{1}{2}N), X_a \leq x \leq N. \end{cases}$$

At this time, the update method of distance accumulation is as follows:

$$D(x, y) = d(x, y) + \min \begin{bmatrix} D(x-1, y) \\ D(x-1, y-1) \\ D(x-1, y-2) \end{bmatrix}. \quad (9)$$

According to the above equation, if the X-axis time label advances gradually, you only need to pay attention to the cumulative distance in the previous column. Therefore, the algorithm does not have to store all the distance matrices, thereby achieving the goal of saving the memory required for the operation. Figure 4 shows the specific update method.

## 5. Experimental Design and Analysis

5.1. *Experimental Design.* To evaluate the performance of the algorithm, you need to collect and analyze the data. In this article, we will select relevant data for mixed education at the

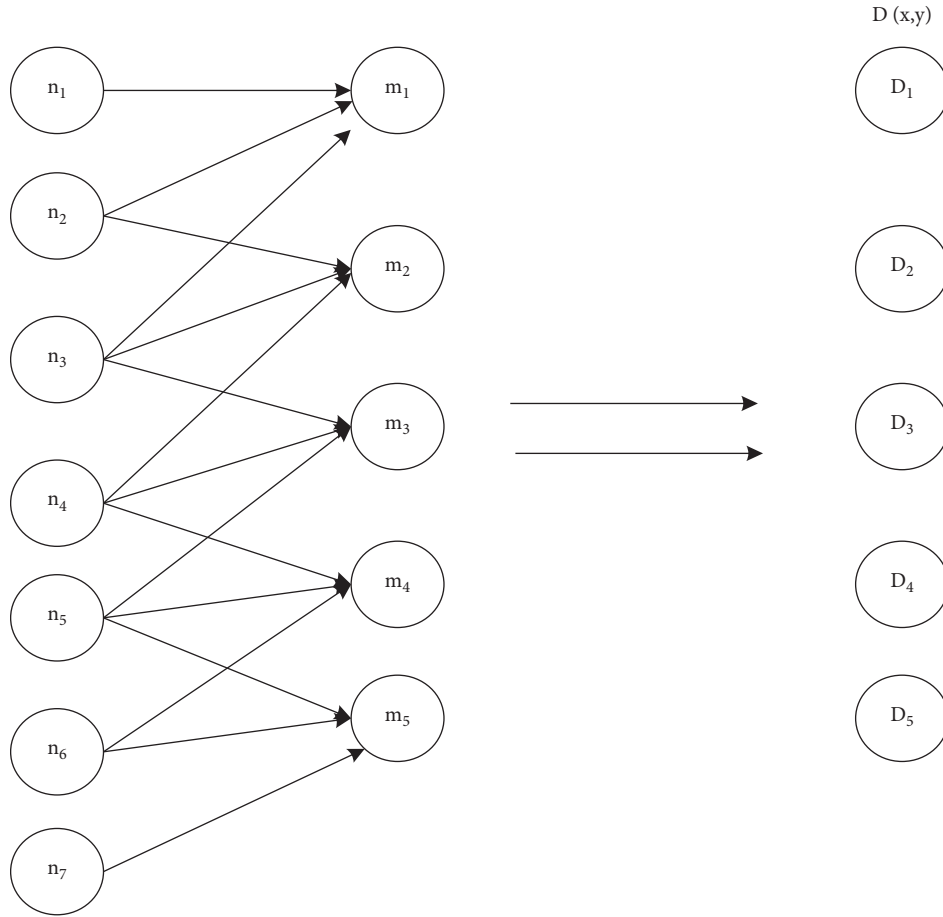


FIGURE 4: Cumulative distance update method.

university. The bibliography [4] method (method 1), bibliography [5] method (method 2), and bibliography [6] method (method 3) are used to test and compare the detection rates of the various methods. The description of the recall rate is as follows: the percentage of anomalous data detected in the overall anomaly of the correct information sample is calculated as follows:

$$\text{detection rate} = \frac{\text{true positive}}{\text{true positive} + \text{false negative}}. \quad (10)$$

In the above formula, true positive represents the true positive rate and true negative represents the true negative class. Recall rate of an abnormal dataset is shown in Figure 5.

From the data in Figure 5, it can be seen that the recall rate of the abnormal points in Method 1 is higher than that in Methods 2 and 3, and the recall rate in Method 1 is as high as 98% or more. 1 is in the diagnostic information for COPD. Prior to the automatic mining of sensitive data, we used an unsupervised step-by-step learning method to clean up sensitive data in COPD diagnostic information. This has reduced a large amount of duplicate data in the database and improved the recall rate. To further verify the overall effectiveness of this method, the method above is used to test the automatic mining of sensitive data for COPD diagnostic information and compare the time spent of the three methods with the same amount of data, Figure 6.

Analyzing Figure 6 shows that with the same amount of data, Method 1 consumes less time than Methods 2 and 3. This is because Method 1 is based on the data cleaning framework for sensitive data in COPD diagnostic information. The cleaning process eliminated erroneous data in COPD diagnostic data, reduced the time required for automatic mining of sensitive data in COPD diagnostic information, and verified the efficiency of Method 1. The error data detection rates of the above three methods are compared. The detection rate of incorrect data can be calculated by the following formula:

$$\text{detection accuracy} = \frac{\text{TN} + \text{TP}}{\text{FN} + \text{TN} + \text{FP} + \text{TP}}. \quad (11)$$

In the above formula, TP represents that the wrong data is judged wrong data, TN represents that the correct data is judged correct data, FP represents that the correct data is judged wrong, and FN represents that the wrong data is judged correct data.

The detection capabilities of the three methods for erroneous data are shown in Figure 7.

As the amount of data increases, the ability to detect incorrect data decreases. Analyzing Figure 7 shows that the error detection rate for Method 1 decreases with increasing data volume, but all exceeds 80%. It automatically resolves and records the optimal boundary function for sensitive data

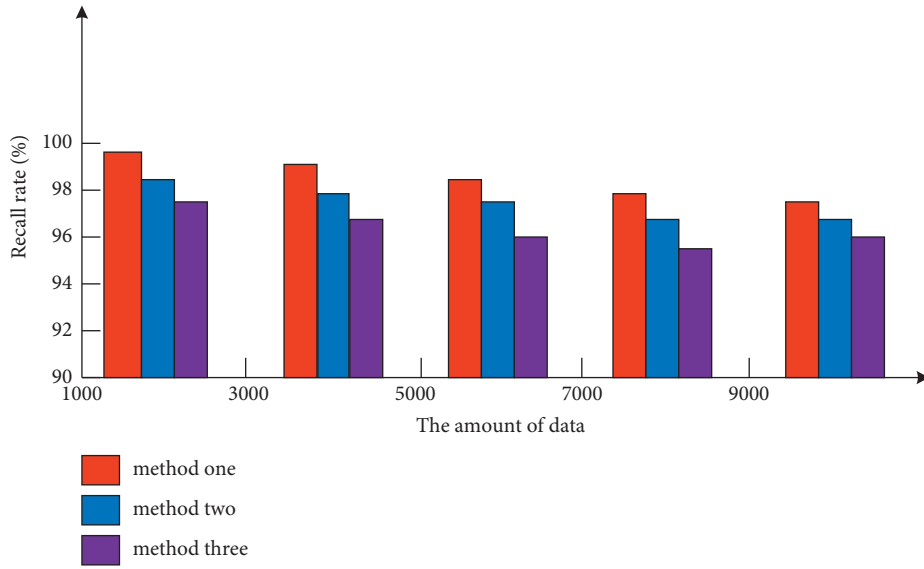


FIGURE 5: Recall rate of abnormal dataset.

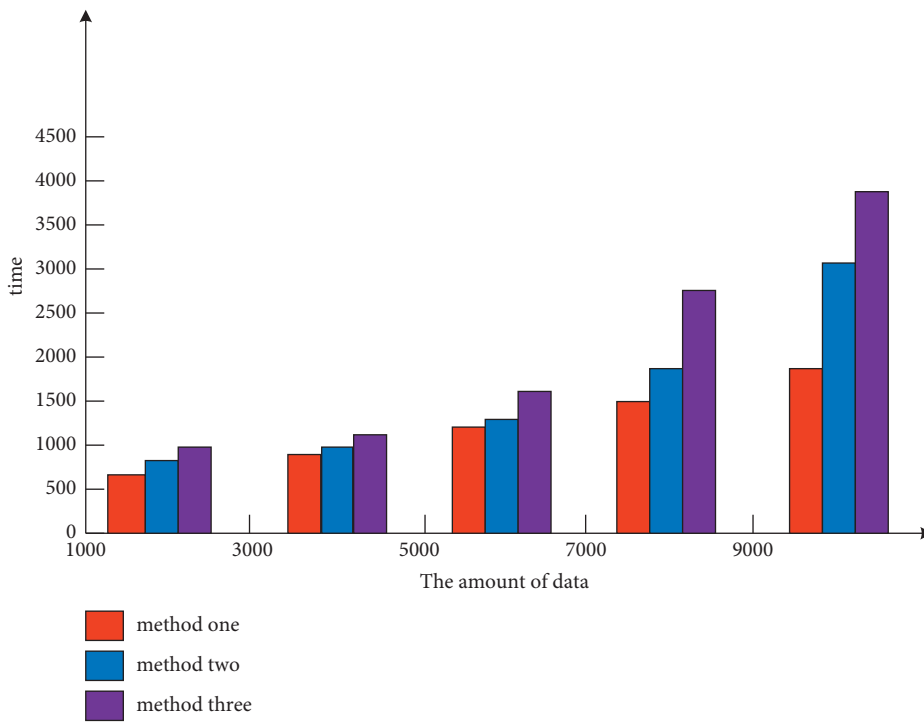


FIGURE 6: Time consumption.

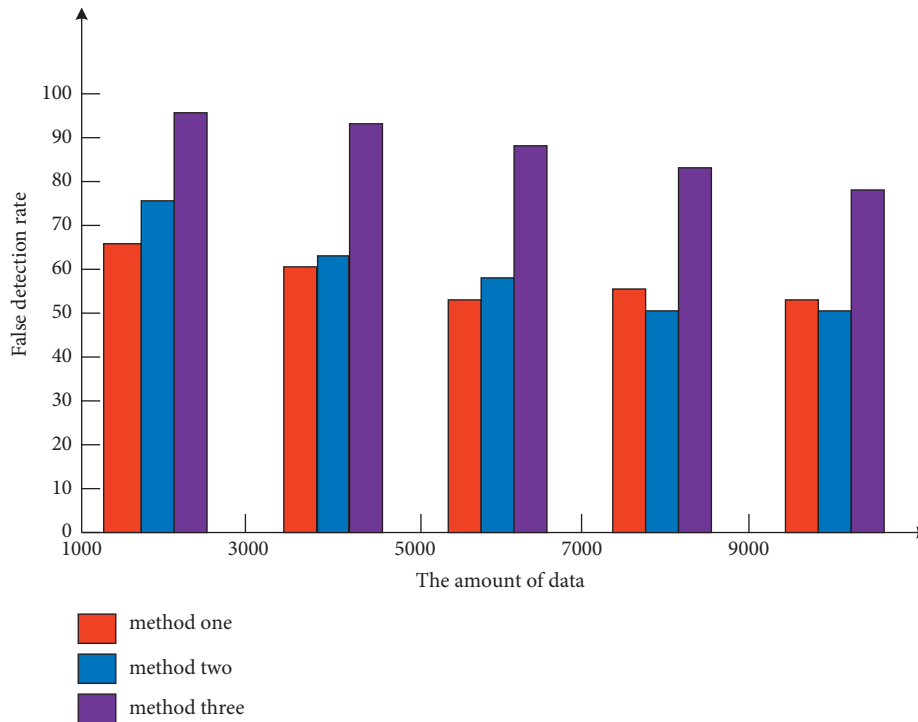


FIGURE 7: Detection rate of wrong data.

in mined COPD diagnostic information, processes large numbers of unlabeled data records, and detects all incorrect data in the database to avoid data nonuniformity. Other useless data improves your ability to detect false data.

## 6. Conclusion

Data mining models can enable the conversion of quantitative data to qualitative assessments and are an effective tool for teaching quality assessments. The data mining model has certain advantages in dealing with mixed educational quality assessments. It can not only get the results of a comprehensive assessment but also reflects the ambiguity and randomness of the assessment data, reflecting the reliability and rationality of the assessment process. In addition, the data mining model method can be used to mine the meaning contained in the evaluation data and obtain more detailed evaluation information so that it is possible to find specific problems in education that are convenient for teachers and improve education in a targeted way, thereby improving educational effectiveness. Of course, when assessing education using a data mining model, there are some issues that make the computational process more complex, and the evaluator must have specific computer application capabilities.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The author declares that he has no conflicts of interest.

## Acknowledgments

This paper was supported in 2017 by the Basic Scientific Research Projects of Colleges and Universities in Liaoning Province “A Practical Study on Microlecture Assisted Teaching Model of College English Course Based on ESP Concept” (Project No. 2017JYT22).

## References

- [1] Y. Zhou, “The application of data mining based on Rough set in teaching evaluation,” *Journal of Chongqing University of Posts and Telecommunications: Natural Science Edition*, vol. 20, no. 5, pp. 627–629, 2008.
- [2] B. Zhang, L. Sun, L. Ma, and H. Xu, “Fuzzy evaluation of university teachers’ teaching quality based on analytic hierarchy process,” *Journal of Liaoning Radio & TV University*, vol. 5, no. 3, pp. 1–4, 2016.
- [3] Y. Zhang, D. Jin, and S. An, “Qualitative evaluation based on cloud model and its application in teaching evaluation,” *Computer Engineering and Applications*, vol. 48, no. 31, pp. 210–215, 2012.
- [4] Y. Ye, “Data mining of classroom teaching quality based on cloud model,” *Computer Simulation*, vol. 29, no. 8, pp. 373–376, 2012.
- [5] C. Liu, F. Mang, X. Dai, and D. Li, “A new reverse cloud algorithm based on cloud X information,” *Journal of System Simulation*, vol. 16, no. 11, pp. 2417–2420, 2004.
- [6] Y. Thoma, A. Dassatti, D. Molla, and E. Petraglio, “FPGA-GPU communicating through PCIe,” *Microprocessors and Microsystems*, vol. 39, no. 7, pp. 565–575, 2015.
- [7] Z. Zhao and C. Qian, “Financial crisis based on data envelopment and data mining Research on computer prediction model,” *Computer Science*, vol. 43, no. S2, pp. 461–465, 2016.

- [8] L. Han and H. Shi, "Application of bayesian classification model to enterprise operation risk prediction," *Microcomputer Applications*, vol. 33, no. 9, pp. 14–10, 2017.
- [9] R. Zhao and W. Changsheng, "Department of enterprise financial management based on hybrid framework system design," *Electronic Design Engineering*, vol. 25, no. 20, pp. 9–12, 2017.
- [10] S. Deng and W. Chen, "Financial data based on a class of support vector machines abnormal pattern recognition," *Journal of Information Engineering University*, vol. 16, no. 2, pp. 251–256, 2015.
- [11] J. Ge, "Research on the design of enterprise financial data management platform," *Electronics Design Engineering*, vol. 24, no. 21, pp. 1–3, 2016.
- [12] J. Sun and L. Qi, "Financial distress prediction model based on GWO-SVMType research," *Microcomputer Applications*, vol. 35, no. 4, pp. 95–99, 2019.
- [13] L. Minmin, H. Jiang, H. Yule et al., "A systematic review on botany, processing, application, phytochemistry and pharmacological action of Radix Rehmanniae," *Journal of Ethnopharmacology*, vol. 6, Article ID 114820, 2021.
- [14] J. Li, Q. Li, and T. Shi, "Wind power based on data miningForecast feature selection method," *Electrical Measurement and Instrumentation*, vol. 56, no. 10, pp. 87–92, 2019.
- [15] Zhanping, Y. Feng, and C. Zhu, "Information analysis based on data mining TechnologyMethod research—taking container shipping price forecast as an example," *Information Science*, vol. 37, no. 7, pp. 65–71, 2019.
- [16] X. Chen, Y. Zhao, and L. Zhang, "Forecast of potential sensitive customers of electric power application of big data method," *Telecommunications Science*, vol. 35, no. 11, pp. 117–124, 2019.
- [17] H. Peng, K. Jin, and C. Fu, "Privacy temporal model based on sequence lattice type mining method," *Acta Electronica Sinica*, vol. 443, no. 1, pp. 156–166, 2020.
- [18] W. Zhijie, T. Jin, and J. Wang, "Medical process differences based on clinical data mining Common discovery methods and applications," *Computer Integrated Manufacturing Systems*, vol. 243, no. 7, pp. 47–57, 2018.
- [19] Z. Fu, B. Liu, and P. Wang, "Coal-fired units based on data mining Energy consumption sensitivity analysis," *Thermal Power Generation*, vol. 47, no. 9, pp. 19–25, 2018.